
PRONÓSTICO DE DEMANDA DE ENERGÍA ELÉCTRICA DE COLOMBIA UTILIZANDO UN MODELO ESTADÍSTICO A PARTIR DE LA METODOLOGÍA DE BOX - JENKINS

Presentado por:

Jeider Augusto Pérez Arenas

Ingeniero Electricista y Mg. en Ingeniería Eléctrica



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES

Facultad de Ingeniería y Ciencias Básicas

Especialización en Estadística Aplicada

Bogotá D.C, Colombia

2020

PRONÓSTICO DE DEMANDA DE ENERGÍA ELÉCTRICA DE COLOMBIA UTILIZANDO UN MODELO ESTADÍSTICO A PARTIR DE LA METODOLOGÍA DE BOX - JENKINS

Presentado por:

JEIDER AUGUSTO PÉREZ ARENAS

en cumplimiento parcial de los requerimientos para optar al título de:

Especialista en Estadística Aplicada

Asesorada por:

Sébastien Lozano Forero

Profesor

LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES

Facultad de Ingeniería y Ciencias Básicas

Especialización en Estadística Aplicada

Bogotá D.C, Colombia

2020

Notas de aceptación



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Firma del presidente del jurado

Firma del jurado

Firma del jurado

Bogotá DC, mayo de 2020.



LOS LIBERTADORES

FUNDACIÓN UNIVERSITARIA

Las directivas de la Fundación Universitaria Los Libertadores, los jurados calificadores y el cuerpo docente no son responsables por los criterios e ideas expuestas en el presente documento. Estos corresponden únicamente a los autores y resultados de su trabajo.

Agradecimientos

En este espacio quiero agradecer a todas las personas que hicieron posible este trabajo, en especial a mi familia por su apoyo incondicional en todo este camino y al profesor Sébastien Lozano Forero por cada momento dedicado para aclarar cualquier tipo de duda que me surgió, agradecerle por la dedicación con la que enseñó cada clase, discurso y lección.

LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Índice general

1	Introducción.....	10
2	Planteamiento del Problema	12
2.1	Objetivos	13
2.1.1	Objetivos General	13
2.1.2	Objetivos Específicos	13
2.2	Justificación.....	14
3	Marco Teórico	15
3.1	Series de Tiempo	15
3.1.1	Metodología Box-Jenkins	17
3.1.2	Criterio BIC y AIC	19
3.1.3	Medición del error	19
3.1.4	Prueba de Dickey-Fuller	20
3.1.5	Modelo ARIMA.....	21
3.1.6	Modelo SARIMA	22
3.1.7	Modelo SARIMAX.....	23
3.2	La demanda de energía eléctrica	23
3.3	Relación entre PIB y IPC.....	24
4	Marco Metodológico.....	26
5	Resultados	28
6	Conclusiones y Recomendaciones	35
7	Referencias Bibliográficas.....	37

Índice de figuras

Figura 1. Demanda anual de energía eléctrica en Colombia vs el PIB(XM S.A. E.S.P., 2015)	11
Figura 2. Proyección de la demanda de energía eléctrica por regiones y total.	12
Figura 3. Metodología Box-Jenkins	17
Figura 4. Demanda de energía eléctrica en Colombia entre el año 2000 y 2018..	24
Figura 5. IPC y Demanda de Energía Eléctrica.....	28
Figura 6. Análisis de tendencia y comportamiento estacional de la demanda de energía eléctrica.....	29
Figura 7. Validación del modelo SARIMAX(2,1;3) x (2; 1; 3) ₁₂ que incluye el cuarto rezago del IPC.....	32
Figura 8. Pronóstico del modelo SARIMAX(2,1;3) x (2; 1; 3) ₁₂ que incluye el cuarto rezago del IPC.....	32
Figura 9. Pronóstico de la demanda de energía eléctrica	33

Índice de tablas

Tabla 1. Parámetros del modelo SARIMAX(2,1;3) x (2; 1; 3) ¹²	30
Tabla 2. Medidas de error del modelo SARIMAX(2,1;3) x (2; 1; 3) ¹²	33

Resumen

Las necesidades de la expansión de la red eléctrica de un país surgen del aumento del consumo de energía eléctrica en el tiempo, cuya proyección se realiza a través de modelos estadísticos con base en información histórica de diferentes variables como el mismo consumo de electricidad, crecimiento demográfico, variables económicas, clima entre otras.

Este trabajo presenta la construcción de un modelo estadístico de tipo SARIMAX, que incorpora como variable exógena el Índice del Precio al Consumidor IPC, generado siguiendo los lineamientos de la metodología de Box -Jenkins. Este tipo de modelo fue elegido, ya que permite incorporar el comportamiento estacional de una serie de tiempo y la inclusión de una variable auxiliar, para este caso el IPC. En la sección 3.3 se explicará como es la relación de esta variable económica con respecto al consumo de electricidad de un país.

La evaluación del modelo fue realizada a través de indicadores de medición del error en unidades de MW y valores porcentuales, encontrando errores inferiores a 90 MW y 1.5% para un periodo de predicción de más de dos años. Estos valores son similares a los errores reportados por la UPME en sus informes. Con el modelo validado fue generado un pronóstico un pronóstico para los 4 meses siguientes a la recolección de los datos.

CAPÍTULO 1

1 Introducción

La electricidad es una de las principales fuentes de energía existentes y es parte fundamental de nuestra vida diaria y pilar de la economía de un país. Con la energía eléctrica no solo alimentamos los electrodomésticos del hogar, sino que también es utilizada como fuente motriz en muchos procesos de la industria. Nos hemos acostumbrado tanto a la energía eléctrica que resulta difícil imaginarnos un mundo sin ella. Es por esto que la carencia de acceso a la energía eléctrica se puede catalogar como símbolo de subdesarrollo económico (Economía.WS, 2018).

Para producir electricidad son utilizadas diferentes fuentes de energía como lo son el agua, el carbón, gas, el sol, el viento entre muchos otros. Y esta energía es transportada grandes distancias mediante las líneas de transmisión, pasando por diferentes niveles de voltaje hasta llegar al consumidor final. En los consumidores finales nos encontramos todos los que hacemos uso de ella, desde el más grande consumidor que son las industrias hasta llegar a nuestros hogares (International Energy Agency, 2017).

En diversos estudios relacionan el crecimiento de la demanda de energía eléctrica con el aumento de la economía del país, aunque no necesariamente en la misma proporción (Galindo Vargas, 2014). Colombia cuenta con un organismo encargado de realizar las predicciones en el consumo de energía eléctrica llamado UPME. La

Unidad de Planeación Minero Energética, es una unidad administrativa especial del orden nacional, de carácter técnico, adscrita al ministerio de minas y energía, cuya principal misión consiste en “*planear de manera integral el desarrollo minero energético, apoyar la formulación de política pública y coordinar la información sectorial con los agentes y partes interesadas*”(«UPME», 2019). En sus informes se puede observar un análisis detallado de cómo será el comportamiento en relación a diversos factores como el Producto Interno Bruto PIB, el aumento en la población y el clima en algunas regiones del país (UPME, 2017), (XM S.A. E.S.P., 2015).

En un informe publicado por el operador del sistema eléctrico colombiano XM, se puede observar las variaciones que ha tenido la demanda de energía eléctrica en el país y la correlación que presenta con respecto al PIB para un periodo de tiempo de 20 años.

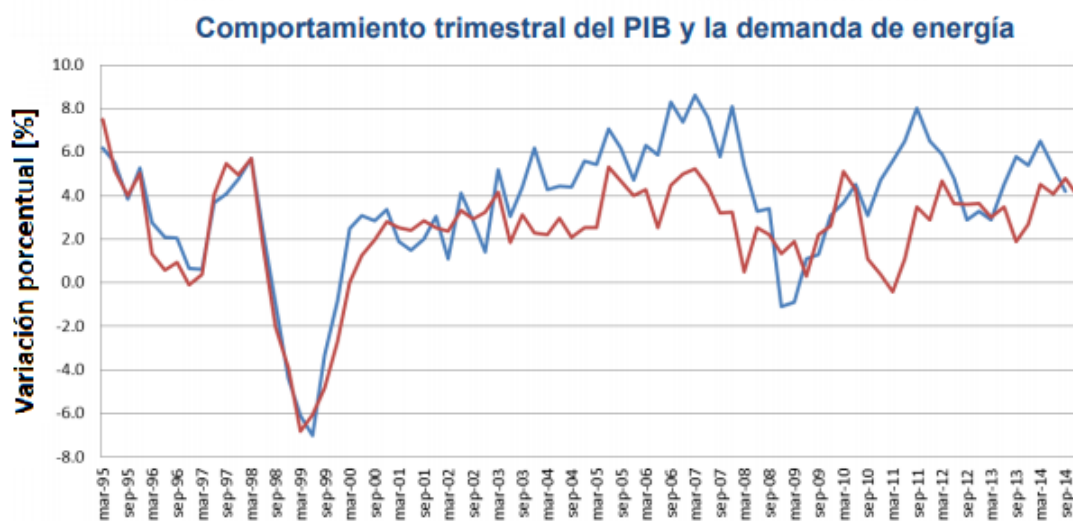


Figura 1. Demanda anual de energía eléctrica en Colombia vs el PIB(XM S.A. E.S.P., 2015)

Hay que tener en cuenta que el PIB es un indicador trimestral y anual. Por lo tanto, no es posible utilizar esta variable exógena en el modelo si se busca un pronóstico mensual, es por eso por lo que en este trabajo se reemplazó el PIB por el IPC. Aunque ambas miden cosas diferentes, se puede decir que existe una estrecha relación entre una y otra.

CAPÍTULO 2.

2 Planteamiento del Problema

La proyección de la demanda de energía eléctrica es de vital importancia para garantizar la confiabilidad de los sistemas eléctricos, ya que con esta es realizada la planeación del sistema para ver las necesidades de proyectos de expansión.

Los escenarios con los que se realiza la expansión en Colombia están ligados a variables externas como métodos econométricos, información regional del clima entre otros.

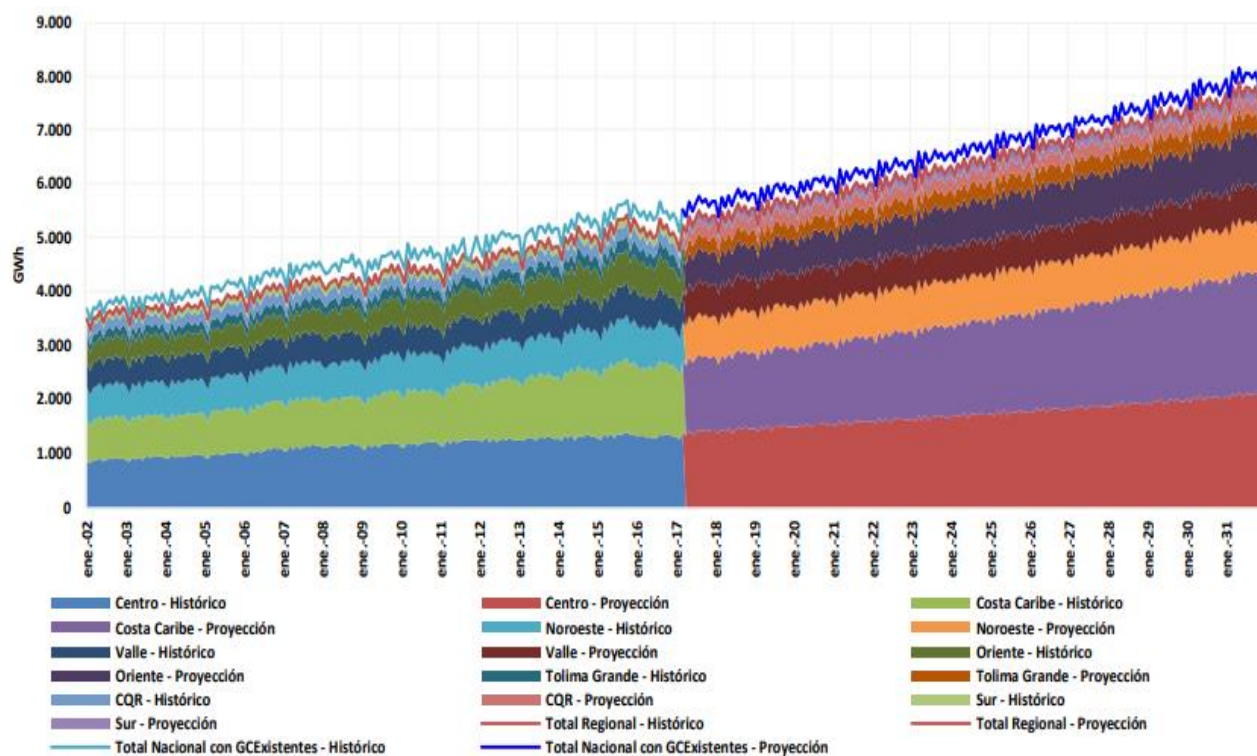


Figura 2. Proyección de la demanda de energía eléctrica por regiones y total.

En la gráfica anterior se presenta la proyección de demanda de energía eléctrica realizado por la UPME para el periodo comprendido entre el año 2018 y el año 2031 distribuidos por regiones. Como se puede observar la proyección fue realizada para un horizonte de 13 años mediante un modelo estadístico que considera tanto variables macroeconómicas como condiciones ambientales (Unidad de Planeación Minero Energética UPME, 2016). Por lo tanto, lo que se busca es un modelo de predicción más sencillo que realice la estimación de la demanda de energía eléctrica para un periodo de tiempo más corto, pero con exactitud.

2.1 Objetivos

2.1.1 Objetivos General

Generar un pronóstico mensual de energía eléctrica en Colombia mediante un modelo estadístico en series de tiempo.

2.1.2 Objetivos Específicos

- Formular un modelo predictivo que permita estimar la demanda de energía eléctrica en Colombia.
- Evaluar y validar el modelo predictivo para estimar la demanda de energía eléctrica en Colombia.
- Con base en el modelo estadístico evaluado y validado crear un pronóstico de demanda de energía eléctrica

2.2 Justificación

Como ya se ha explicado la importancia de proyectar la demanda de energía eléctrica, ya que los proyectos de expansión de la red eléctrica se hacen con base en las necesidades que se reflejan a través de los pronósticos de consumos. Adicionalmente, la proyección de la demanda es utilizada para realizar subastas de energía, con el fin de garantizar la confiabilidad en el suministro a precios eficientes e incentivar proyectos de generación de energía eléctrica (Unidad de Planeación Minero Energética UPME, 2016).

El modelo que considera la UPME es un modelo bastante complejo compuesto por diferentes variables como lo son el comportamiento de la economía a través del PIB, el crecimiento de la población y factores ambientales como la temperatura. (Unidad de Planeación Minero Energética UPME, 2016). Lo que se busca en este trabajo es desarrollar un modelo estadístico haciendo uso de las series de tiempo que sirva para pronosticar la demanda de energía eléctrica en Colombia utilizando el Índice del Precio al Consumidor IPC como variable externa.

CAPÍTULO 3

3 Marco Teórico

3.1 Series de Tiempo

Una serie tiempo X_t representa una secuencia de observaciones medidas uniformemente en el tiempo y ordenadas cronológicamente, las cuales presentan una dependencia entre sí. Una serie de tiempo se puede definir como la unión de tres componentes tal como se representa en la siguiente ecuación(Rios & Hurtador, 2008):

$$X_t = T_t + E_t + I_t$$

Donde:

T_t : Componente tendencia.

E_t : Componente estacional.

I_t : Componente aleatoria.

- Componente tendencia: Se define como el cambio a largo plazo en el valor de la media.

- **Componente estacional:** Muchas series temporales presentan cierta periodicidad en las observaciones, puede ser semestral, mensual, etc. Estos efectos son fáciles de entender y se pueden medir explícitamente o incluso se pueden eliminar de la serie de datos, a este proceso se le llama desestacionalización de la serie.
- **Componente aleatoria:** Esta componente no responde a ningún patrón de comportamiento, sino que es el resultado de factores fortuitos o aleatorios que inciden de forma aislada en una serie de tiempo.

Las series se pueden clasificar en:

- **Estacionarias:** Una serie es estacionaria cuando su varianza y su media son constantes en el tiempo. Visualmente se puede determinar observando si la serie oscila alrededor de una media constante y la variabilidad con respecto a ese valor también lo es.
- **No estacionarias:** Una serie es no estacionaria cuando la tendencia y/o variabilidad cambian en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante.

Para que una serie de tiempo sea estacionaria se debe cumplir las siguientes propiedades:

$$\text{Media } E(X_t) = E(X_{t+k}) = \mu$$

$$\text{Varianza } V(X_t) = V(X_{t+k}) = \sigma^2$$

$$\text{Covarianza } \gamma_k = E[(X_t - \mu)(X_{t+k} - \mu)]$$

Donde γ_k es la covarianza al rezago k . En resumen, si una serie de tiempo es estacionaria, su media, su varianza y su covarianza en diferentes rezagos permanecen iguales sin importar el momento en el cual se midan; es decir, son invariantes respecto al tiempo.

3.1.1 Metodología Box-Jenkins

Es una metodología que surgió en los años 70's, también es conocida como metodología ARIMA. Esta metodología busca estimar un modelo estadístico que sirva para realizar predicciones futuras basados en los valores pasados de un conjunto de datos. Se divide en 4 etapas descritas a continuación. (Rosales, Urrego, Sánchez, Correa, & Morales, 2008)

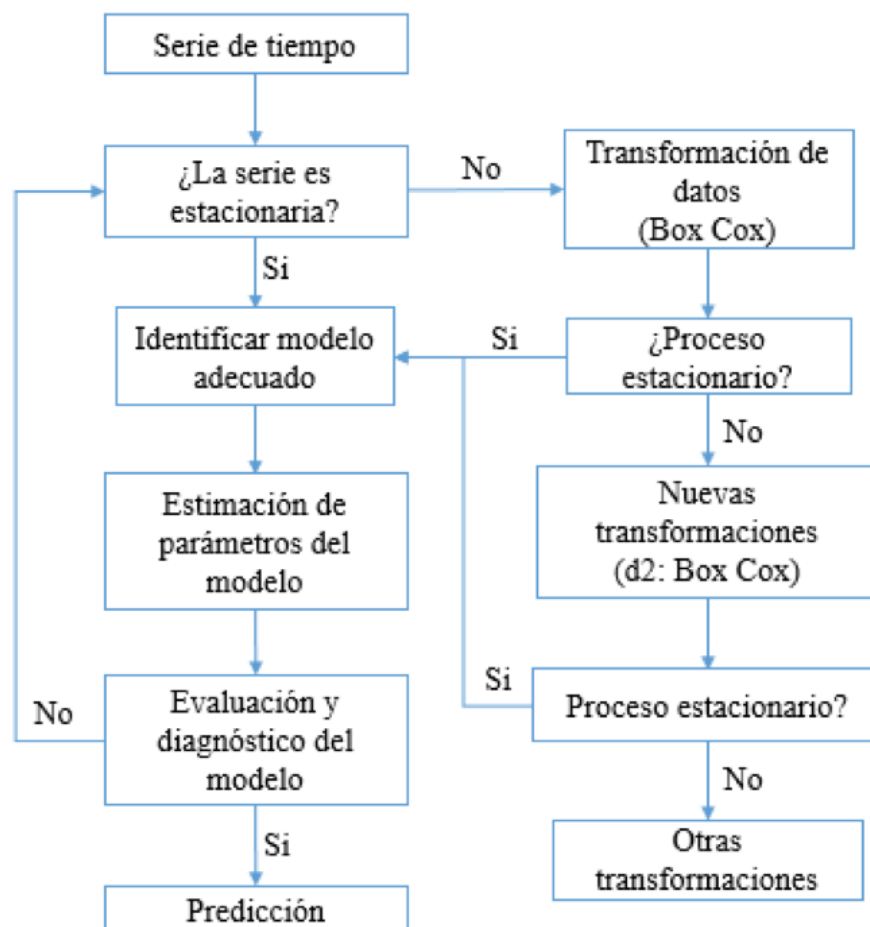


Figura 3. Metodología Box-Jenkins

- **Etapla 1: Identificación**

En esta etapa a partir de los coeficientes de correlación simple y parcial se busca identificar el tipo de modelo que mejor se adapte a los datos. Es importante tener en cuenta que lo primero que se debe hacer es verificar que la serie de tiempo sea estacionaria

- **Etapla 2: Estimación**

En esta etapa se estiman los coeficientes de los términos autorregresivos y de media móvil incluidos en el modelo, cuyo número de rezagos han sido identificados en la etapa anterior.

- **Etapla 3: Verificación**

Esta etapa busca evaluar si el modelo estimado se ajusta a los datos en forma razonablemente buena, para esto se utilizan diferentes pruebas entre las cuales se encuentran.

- **Análisis de los coeficientes:** El modelo estimado debe cumplir con condiciones de estacionariedad e invertibilidad y que los rezagos incorporados deben ser estadísticamente significantes.
- **Bondad del ajuste:** A partir de indicadores determinar qué modelo se ajusta mejor a los datos. Entre los indicadores se encuentran el R^2 Ajustado, AIC y BIC.
- **Análisis de los residuos:** Se parte del supuesto que los errores del modelo son un proceso aleatorio (media cero, varianza constante y no correlación serial).

Si el modelo no cumple con la etapa de verificación, se volverá a la etapa de identificación con el objeto de reformular el modelo

- **Etapla 4: Pronóstico**

Utilizar el modelo para pronosticar un periodo futuro.

3.1.2 Criterio BIC y AIC

Cuando se tiene una serie de modelos M_1, M_2, \dots con parámetros K_1, K_2, \dots , respectivamente, se pueden comparar mediante medidas de contraste como el AIC y el BIC. Ambos criterios hacen uso del Log-likelihood ($\log Lik$), que es el logaritmo de máxima verosimilitud y sustraen un término proporcional al número de parámetros (K) en el modelo (Posada & Noguera, 2007).

- **Criterio de información de Akaike (AIC):**

Este criterio tiene en cuenta los cambios en la bondad de ajuste y las diferencias en el número de parámetros entre dos modelos. Los mejores modelos son aquellos que presentaron el menor valor de AIC.

$$AIC = -2 * \log Lik + 2K$$

- **Criterio de información bayesiano (BIC):**

El BIC es calculado para los diferentes modelos como una función de la bondad de ajuste del $\log Lik$, el número de parámetros ajustados (K) y el número total de datos (N). El modelo con el más bajo valor de BIC es considerado el mejor en explicar los datos con el mínimo número de parámetros.

$$BIC = -2 * \log Lik + \log(N) * K$$

3.1.3 Medición del error

La evaluación de un modelo se cuantifica a partir de qué tan cerca está la variable pronosticada de su serie de datos correspondiente. A continuación, se muestran una serie de indicadores cuyo objetivo es comparar que tan cerca están los n pronosticados con sus respectivos valores reales:

Considerando:

$$e_j = y_j - \hat{y}_j$$

Donde \hat{y}_j denota el pronóstico de y_j

- Error medio (ME):

$$ME = \frac{1}{n} \sum_{j=1}^n e_j$$

- Raíz de error cuadrático medio (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n e_j^2}$$

- Error absoluto medio (MAE):

$$MAE = \frac{1}{n} \sum_{j=1}^n |e_j|$$

- Error medio porcentual (MPE):

$$MAE = \left(\frac{1}{n} \sum_{j=1}^n \frac{e_j}{y_j} \right) \times 100$$

- Error Absoluto medio porcentual (MAPE):

$$MAPE = \left(\frac{1}{n} \sum_{j=1}^n \frac{|e_j|}{y_j} \right) \times 100$$

3.1.4 Prueba de Dickey-Fuller

Al realizar el análisis de series de tiempo, la mayoría de los métodos de pronóstico estadístico asumen que las series de tiempo son aproximadamente estacionarias. Se define como serie estacionaria aquellas cuyas propiedades estadísticas como la

varianza y la media son constantes en el tiempo. Una de las formas de determinar si la serie es estacionaria es la prueba de Dickey-Fuller y Dickey-Fuller Aumentada.

Suponiendo que la variable sigue un modelo AR(1) se tiene:

$$y_t = \phi y_{t-1} + u_t$$

Aplicando el operador de rezagos L

$$y_t - \phi y_{t-1} = u_t \Rightarrow y_t - \phi L y_t = u_t \Rightarrow (1 - \phi L) y_t = u_t$$

La solución es $y_t = \lambda^t$, donde λ es la raíz característica, siendo $\lambda = \phi$ y se analizan dos posibles soluciones.

- $\lambda < 1 \Rightarrow \lim_{t \rightarrow \infty} \phi^t = 0$. Por lo tanto, la esperanza $E(y_t) = 1/(1 - \phi)$
- $\lambda \geq 1 \Rightarrow \lim_{t \rightarrow \infty} \phi^t = \infty$. Por lo tanto, la esperanza $E(y_t) = \infty$.

La existencia de una raíz unitaria hace que el valor esperado de la serie sea infinito. Suponiendo que los datos siguen un modelo AR(p), entonces aparece la prueba de y Dickey-Fuller Aumentada la cual utiliza el siguiente modelo:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^p \lambda_i \Delta y_{t-1} + u_t$$

El modelo anida las diferentes alternativas de prueba, es decir se puede efectuar la prueba de raíz unitaria con o sin constante, con o sin tendencia determinística y con o sin considerar autocorrelaciones.

3.1.5 Modelo ARIMA

Los modelos ARIMA (“Auto Regressive Integrated Moving Average”) resultan de la integración de dos tipos de modelos, los auto regresivos AR y los de media móvil MA

desarrollado por Box-Jenkins, los cuales se basan en el tratamiento de la correlación de la serie(Lozano Forero, 2018).

Se parte del hecho de que la serie no es estacionaria, pero $\Delta^d y_t = (1 - B)^d y_t$ si lo es. La ecuación general de un modelo ARIMA se define como:

$$\phi(B)(1 - B)^d y_t = \theta(B)\epsilon_t$$

Donde:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad \text{Polinomio auto-regresivo}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad \text{Polinomio de medias móviles}$$

3.1.6 Modelo SARIMA

Los modelos SARIMA captan el comportamiento estacional de una serie de tiempo ARIMA(P,D,Q), asociada a observaciones que se encuentran separadas s periodos. Si se adiciona la componente ARIMA(p,d,q) que modela la componente no estacional o regular la cual está asociada a observaciones consecutivas se llega al modelo SARIMA multiplicativo(Lozano Forero, 2018).

Sea el modelo de la parte estacional:

$$\Phi(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_P B^{sP} \quad \text{Operador regresivo estacional.}$$

$$\Theta(B^s) = 1 - \theta_1 B^s - \dots - \theta_Q B^{sQ} \quad \text{Operador de media móvil estacional.}$$

El modelo multiplicativo (p; d; q) x (P;D;Q) está dada por:

$$\begin{aligned} & (1 - \phi_1 B - \dots - \phi_p B^p)(1 - \Phi_1 B^s - \dots - \Phi_P B^{sP})[(1 - B)^d(1 - B^s)^D]y_t \\ & = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)(1 - \Theta_1 B^s - \dots - \Theta_Q B^{sQ})\epsilon_t \end{aligned}$$

Donde $\epsilon_t \sim RB(0, \sigma^2)$. Reescribiendo el modelo de forma simplificada se tiene que:

$$\phi(B)\Phi(B^s)(1 - B)^d(1 - B^s)^D y_t = \theta(B)\Theta(B^s)\epsilon_t$$

3.1.7 Modelo SARIMAX

Un modelo SARIMAX es la combinación de un modelo SARIMA con la propiedad que permite incorporar en el modelamiento variables exógenas que pueden ayudar a entender y pronosticar la variable de interés(Lozano Forero, 2018).

Para el caso en que se tiene una única variable de interés y , y una única variable independiente x . El modelo que describe la relación entre y en un tiempo t en términos de r observaciones anteriores de x está dado por:

$$y_t = \alpha + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + B_0 x_t + B_1 x_{t-1} + \dots + B_r x_{t-r} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}$$

Donde α , B_i , ϕ_j , θ_k , con $i = 0, \dots, r$. $j = 0, \dots, p$. $k = 0, \dots, q$. Son parámetros fijados y $\epsilon_t \sim RB(0, \sigma^2)$.

De forma general el modelo puede ser escrito como:

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D y_t = \theta(B)\Theta(B^s)\epsilon_t + \sum_{j=1}^r B_j x_{t-j}$$

3.2 La demanda de energía eléctrica

El consumo de energía eléctrica está asociado a la cantidad de energía eléctrica consumida por todos los equipos eléctricos y electrónicos en un periodo de tiempo determinado. Se ve afectada por factores como el clima, en los días fríos existe una disminución en el consumo debido a la baja utilización de aires acondicionados, en épocas de vacaciones y decembrinas muchas industrias suspenden labores por lo que también se ve reflejada una disminución en el consumo de energía eléctrica y así existen muchos factores más como económicos, en los cuales un aumento en la economía del país indica implícitamente que la industria está generando mayor producción por lo que finalmente se ve reflejado en un aumento de la energía eléctrica.

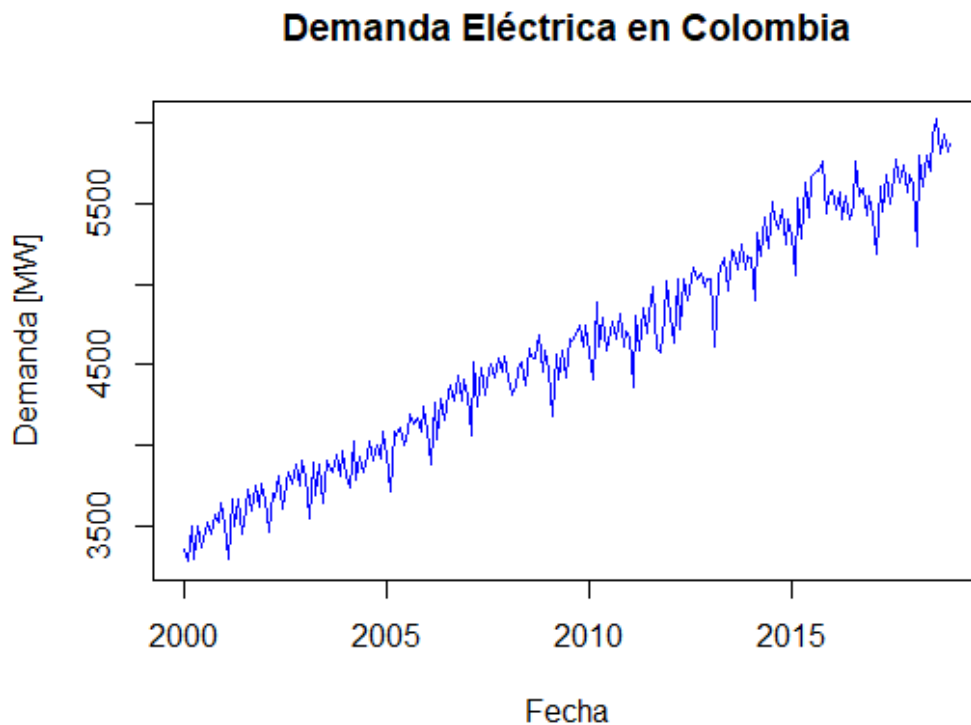


Figura 4. Demanda de energía eléctrica en Colombia entre el año 2000 y 2018

De la Figura 4 se puede observar como el consumo de energía eléctrica en Colombia en el periodo de tiempo mostrado presenta una tendencia creciente, aumentando en este intervalo de tiempo en más del 50% del consumo en el país. Adicionalmente se puede apreciar un comportamiento estacional en cada año, por lo que el modelo que se propone es un SARIMAX descrito en la sección 3.1.7, donde se buscará un indicador económico como variable regresora.

3.3 Relación entre PIB y IPC

Por definición el Producto Interno Bruto (PIB) es el valor de mercado de todos los bienes y servicios finales producidos dentro de un país, en un periodo determinado. Por lo general, se determina mediante el consumo, la inversión, compras del gobierno y exportaciones netas, generalmente se calcula de forma trimestral y un

acumulado anual. El Índice de Precio al Consumidor (IPC) es una medida de los precios agregados y se calcula como una media ponderada de los precios de una canasta de bienes y servicios finales representativo del consumo de los hogares (Casares Hontañón & Tezanos Vázquez, 2009).

Aunque son dos variables que miden dos cosas diferentes, mantienen una relación estrecha entre sí. Se dice que, si el PIB crece, produce un aumento en el IPC. Si el crecimiento ocurre demasiado rápido, la inflación también crece demasiado rápido, por lo que el costo de vida se vuelve demasiado alto como para que las personas puedan ajustarse a él. Adicionalmente las personas no pueden costear los nuevos precios de los productos debido a que el aumento de los ingresos es más lento que la inflación (Casares Hontañón & Tezanos Vázquez, 2009).

CAPÍTULO 4.

4 Marco Metodológico

Este informe busca encontrar un modelo estadístico para la predicción de la demanda de energía eléctrica en Colombia. La información utilizada en ese trabajo es de acceso público y fue tomada de dos fuentes diferentes, el IPC fue tomado del portal web del Departamento Administrativo Nacional de Estadística (DANE, 2019) y los valores de demanda de energía eléctrica del portal BI de la empresa operadora y administradora del sistema eléctrico colombiano (XM S.A. E.S.P., 2019)

En base a la revisión bibliográfica y al marco teórico realizado, se determinó que la mejor alternativa para modelar el comportamiento de la demanda de energía eléctrica en el tiempo es un modelo SARIMAX (Auto Regresivo Integrado de Medias Móviles Estacional con variables exógenas). Ya que este modelo permite integrar los datos de la demanda eléctrica con el Índice del Precio al Consumidor (IPC).

La metodología para ajustar el modelo está basada en la metodología de Box-Jenkins descrita en la sección anterior. En primera instancia se realizó una búsqueda de los tipos de modelos que se pudieran ajustar a los datos empleados determinando que la mejor opción es un modelo SARIMAX.

Posteriormente se realizó un proceso iterativo buscando encontrar el mejor ajuste de los parámetros del modelo SARIMAX significativos mediante el valor-t y evaluado mediante las gráficas de autocorrelación ACF y PACF y el índice BIC, resultando en un modelo de parámetros SARIMAX(2,1;3) x (2; 1; 3)¹².

Con este modelo el siguiente paso consistió en evaluar la calidad del pronóstico midiendo el error de estimación descrito en la sección 3.1.3, evaluando la normalidad de los residuos y verificando que no exista correlación serial entre los mismos. Finalmente, como resultado final se realiza un pronóstico de la demanda de energía eléctrica para el sistema colombiano.

CAPÍTULO 5.

5 Resultados

En esta sección se presenta la creación y validación del modelo estadístico creado a partir de los datos de demanda de energía eléctrica y del Índice del Precio al Consumidor IPC mostrados en la Figura 5.

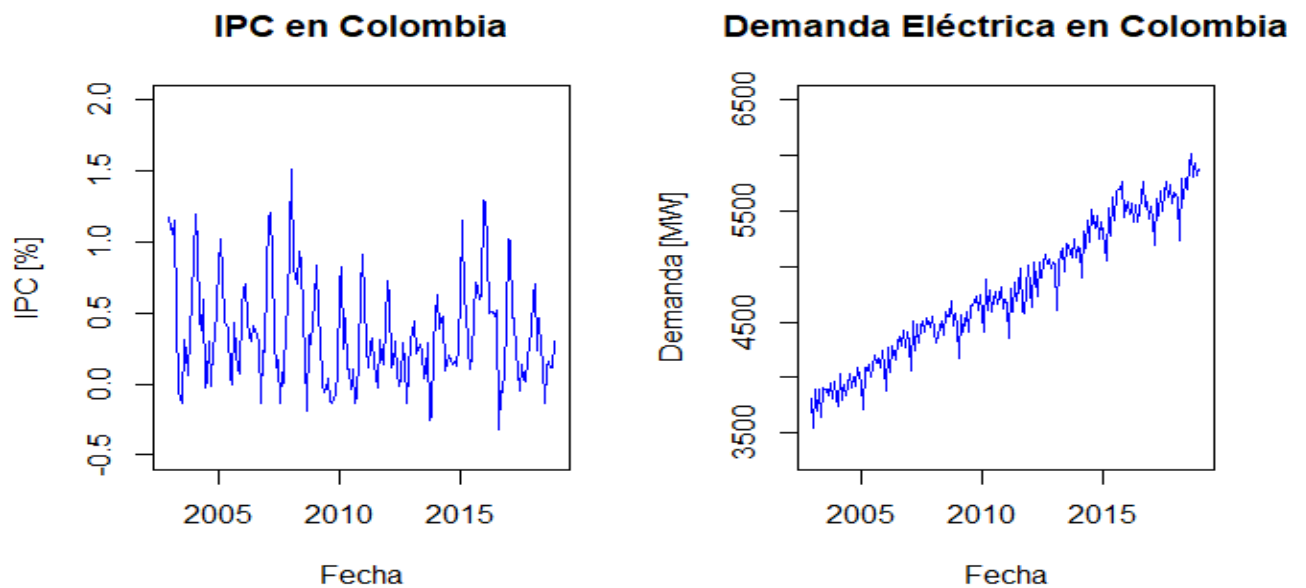


Figura 5. IPC y Demanda de Energía Eléctrica

La metodología empleada para la elaboración del modelo y su validación fue descrita en el capítulo 4. En resumen, se recopiló información del IPC y de la demanda de energía eléctrica entre enero de 2013 y junio de 2019. Estos datos fueron divididos en dos grupos, los primeros comprendidos entre enero de 2013 y diciembre de 2016 fueron utilizados para crear el modelo estadístico llamándose como los datos dentro de la muestra (*in-sample*). El segundo grupo de datos comprendidos entre enero de 2017 y junio de 2019 corresponden a los datos de validación del modelo y se les conoce como los datos fuera de la muestra (*out-sample*), es decir el modelo deberá ser capaz de realizar un buen pronóstico de los valores de demanda de energía eléctrica por fuera de los datos con los cuales fue realizado.

En un análisis preliminar realizado al comportamiento histórico de la demanda de energía eléctrica, se observa una tendencia creciente en el tiempo la cual se ajusta a un modelo de regresión lineal, esta tendencia se puede eliminar realizando la diferenciación a la serie tal como se presenta en la Figura 6 donde se observa que el promedio y la varianza de la serie se estabilizan.

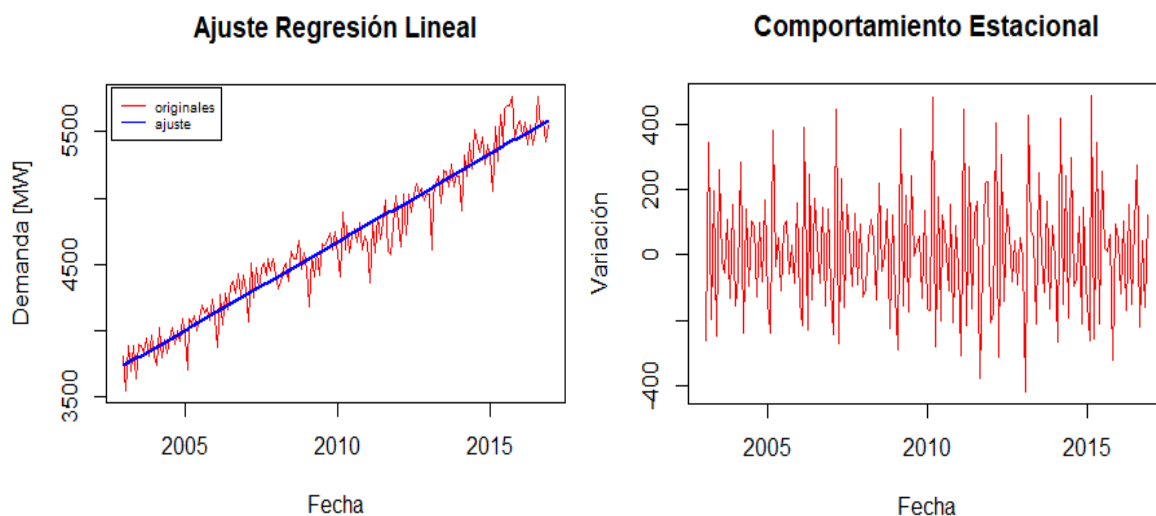


Figura 6. Análisis de tendencia y comportamiento estacional de la demanda de energía eléctrica.

La estacionalidad de la demanda de energía eléctrica la caracteriza el consumo típico mensual, definido en parte por un comportamiento climático, un comportamiento cultural y económico, el cual tiende a ser repetitivo en el tiempo.

Luego de un proceso iterativo se llega a un modelo de tipo SARIMAX (2,1;3) x (2; 1; 3)₁₂ definidos por los parámetros descritos en la Tabla 1 y con un valor de AIC y BIC de 1830 y 1855 respectivamente. Los parámetros que no están contenidos en esta tabla se consideran en cero debido a que su valor t no es significativo (inferior a 1.90).

Parámetro	ar ₁	ar ₂	ma ₃	sar ₁	sar ₂	sma ₃	IPC ₄
Estimación	-0.65	-0.49	-0.26	-1.01	-0.99	-1.0	-52.09
Error Estandár	0.08	0.097	0.11	0.02	0.05	0.19	27.33
t - valor	8.16	5.07	2.43	44.20	20.64	5.25	1.91

Tabla 1. Parámetros del modelo SARIMAX(2,1;3) x (2; 1; 3)₁₂

La ecuación general del modelo fue descrita en la sección 3.1.7. Donde se tiene:

$$\phi_p(B)\Phi_p(B^s)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t + \sum_{j=1}^r B_j x_{t-j}$$

Descomponiendo cada término de la ecuación (Izquierda) y reemplazando por los valores del modelo (Derecha):

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rightarrow \phi_p(B) = (1 + 0.65B + 0.49B^2)$$

$$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \rightarrow \theta_q(B) = (1 + 0.26B^3)$$

$$\Phi_P(B^s) = (1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}) \rightarrow \Phi_P(B^s) = (1 + 1.01B^{12} + 0.99B^{24})$$

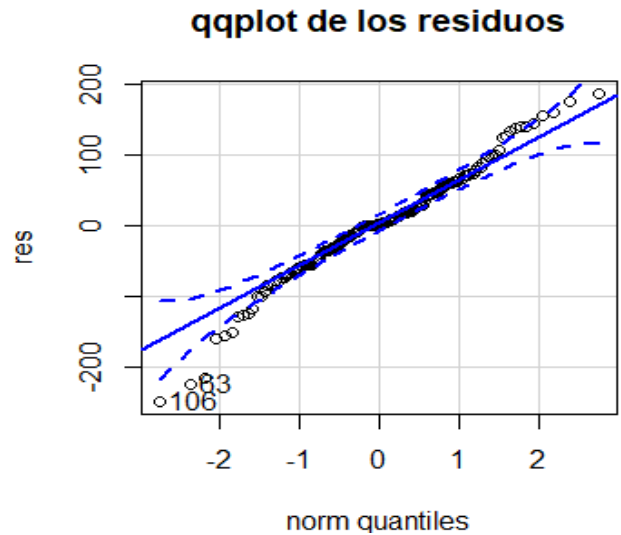
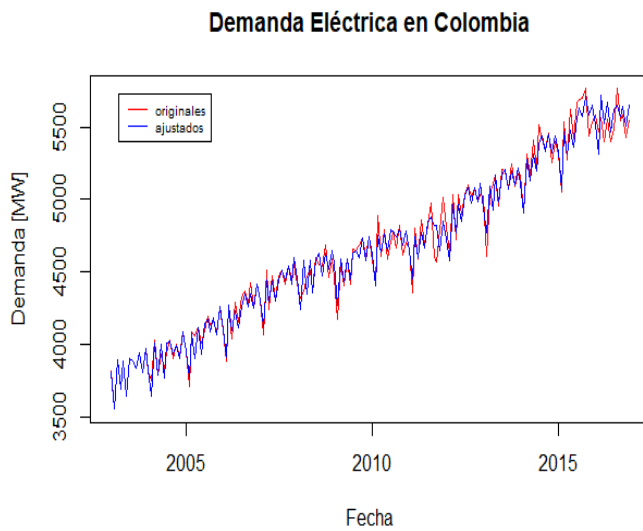
$$\Theta_Q(B^s) = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}) \rightarrow \Theta_Q(B^s) = (1 + B^{36})$$

$$\sum_{j=1}^r B_j x_{t-j} \rightarrow -52.09 IPC_{t-4}$$

$$d \rightarrow 1$$

$$D \rightarrow 1$$

La Figura 7 está compuesta de 4 gráficas utilizadas para el diagnóstico del modelo. La primera consiste en un comparativo de las demanda de energía eléctrica real con la ajustada por el modelo para el periodo entre enero de 2003 y diciembre de 2016. Seguido se encuentra un gráfico de Cuantil – Cuantil (qqplot) de los residuos, en el cual se observa que estos no siguen una distribución normal y se confirma a través del p-valor mediante la prueba de Jarque-Bera, obteniendo un valor de $1,29 \times 10^{-3}$, sugiriendo con una significancia del 5% que los residuos no se distribuyen normalmente y finalmente los gráficos de autocorrelación y autocorrelación parcial de los residuos donde no se evidencia ningún problema de autocorrelación serial y se confirma mediante el test de Ljung-Box obteniendo un valor de 0.76 para 60 lazos.



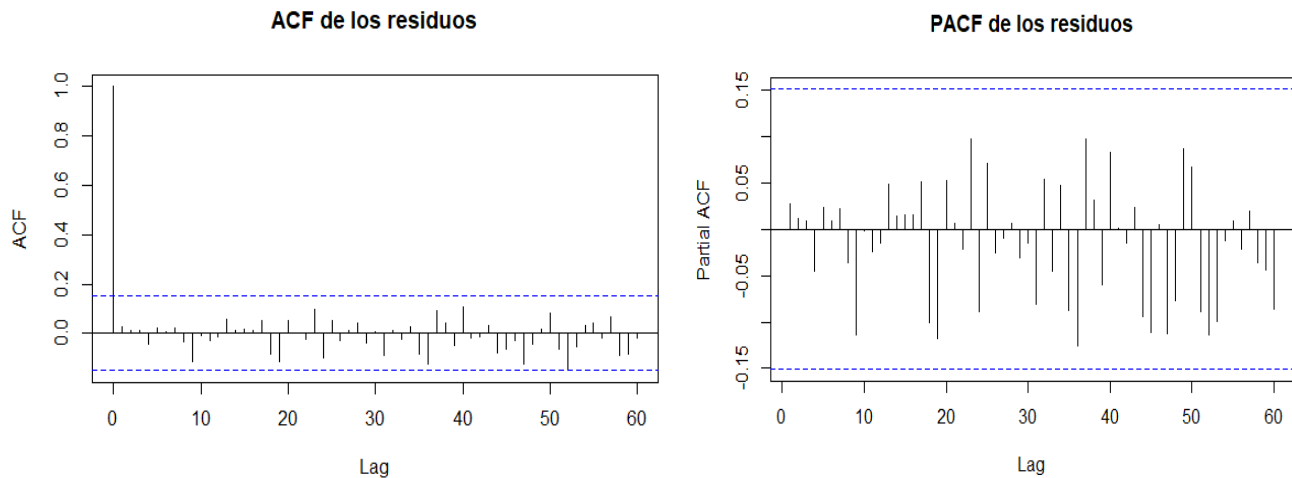


Figura 7. Validación del modelo $SARIMAX(2,1;3) \times (2; 1; 3)_{12}$ que incluye el cuarto rezago del IPC

Una vez verificada las condiciones con las cuales el modelo fue realizado, se debe verificar la calidad de la estimación con los datos fuera de la muestra (*out-sample*) y los índices de medición del error descritos en la sección de 3.1.3. En la Figura 8 se presenta un comparativo entre el pronóstico generado con el modelo en color azul y los valores reales en color rojo para enero de 2017 hasta junio de 2019. Adicionalmente en una franja gris se presenta el intervalo de confianza del 95% del pronóstico, donde se observa que el pronóstico generado no sobrepasa esta franja en ningún punto de la curva.

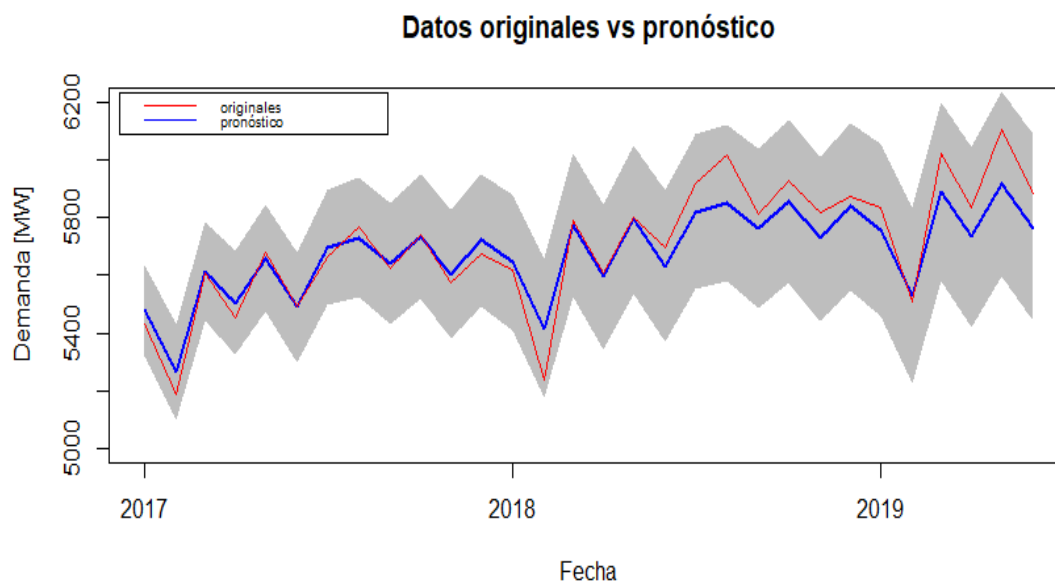


Figura 8. Pronóstico del modelo $SARIMAX(2,1;3) \times (2; 1; 3)_{12}$ que incluye el cuarto rezago del IPC

En la siguiente tabla se presenta los resultados del error de estimación del pronóstico en el proceso de validación del modelo. Los tres primeros índices se encuentran en unidades de [MW] y hacen referencia al error esperado en cada estimación. Sin embargo, los índices más dicientes son el MPE y MAPE cuyos valores se encuentran en porcentaje, por tal razón es mucho más fácil su interpretación. Para este modelo ambos valores no superan el 2%, por lo que se puede afirmar una buena calidad en la estimación.

ME [MW]	RMSE [MW]	MAE [MW]	MPE [%]	MAPE [%]
35.67	84.35	64.62	0.58	1.12

Tabla 2. Medidas de error del modelo SARIMAX(2,1;3) x (2; 1; 3)¹²

Con el modelo descrito en la Tabla 1 se realiza el pronóstico de la demanda de energía eléctrica mostrado en la Figura 9 para el periodo entre julio de 2019 y octubre de 2019.

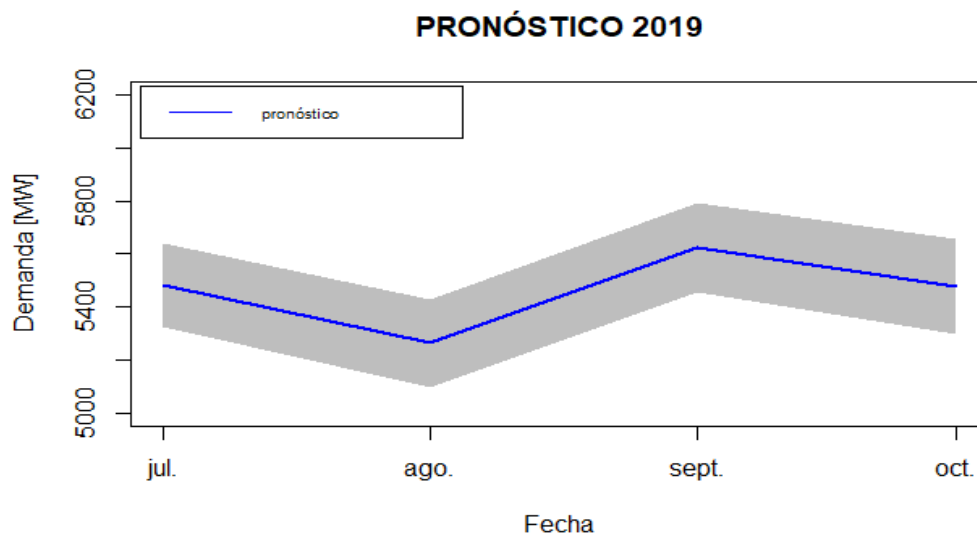


Figura 9. Pronóstico de la demanda de energía eléctrica

El modelo presentado toma como variable regresora la información del IPC en el cuarto rezago. Es decir, para conocer el valor de la demanda de energía eléctrica en el mes k debemos conocer el valor del IPC en el mes $k-4$. Ya que a la fecha de realización de este trabajo solo se tiene información del IPC hasta el mes de junio de 2019, el pronóstico máximo obtenido es el mes de octubre de 2019.

Además de lo expuesto en la justificación del trabajo, este pronóstico servirá a futuro para realizar la comparación de la demanda de energía eléctrica realizada por el modelo con los valores reales y verificar la calidad del pronóstico.

CAPÍTULO 6.

6 Conclusiones y Recomendaciones

- Haciendo uso de la información histórica de la demanda de energía eléctrica y el índice del precio al consumidor IPC en Colombia para un periodo comprendido entre el año 2003 y el año 2016 se logró desarrollar un modelo estadístico de tipo SARIMAX mediante la metodología de Box-Jenkins, con el cual se realizó el pronóstico de la demanda de energía eléctrica para los meses julio a octubre de 2019.
- Se realizó la evaluación y validación del modelo estadístico mediante el uso de indicadores de medición del error y de ajuste. Estos indicadores sirvieron como referente para escoger el mejor modelo dentro del proceso iterativo realizado a través de la metodología de Box-Jenkins.
- Se determinó que las medidas de error MPE y MAPE al ser porcentuales funcionan mejor en el proceso de evaluación de la calidad de pronóstico del modelo, obteniendo errores por debajo del 2% para ambas mediciones, lo cual es un buen valor de tolerancia de estimación.

- Se recomienda investigar nuevas relaciones de variables macroeconómicas con el consumo de energía eléctrica en Colombia, que permitan mejorar el ajuste del modelo estadístico generado.
- Como trabajo futuro se plasma realizar un comparativo en la generación de pronósticos mensuales de energía eléctrica con diferentes tipos de modelos estadísticos, esto permitirá determinar las ventajas y desventajas de cada modelo y encontrar el que mejor se ajuste al problema.

7 Referencias Bibliográficas.

Casares Hontañón, P., & Tezanos Vázquez, S. (2009). *Principios de Economía*.

Recuperado de

<https://ocw.unican.es/pluginfile.php/977/course/section/1110/Principios-economia5.pdf>

DANE. (2019, julio 29). Índice de Precios al Consumidor IPC. Recuperado 29 de julio de 2019, de <https://www.dane.gov.co/>

Economía.WS. (2018). Energía eléctrica—Economía.WS. Recuperado 20 de junio de 2018, de <http://www.economia.ws/energia-electrica.php>

Galindo Vargas, A. P. (2014). *La relación entre el consumo de electricidad y el crecimiento económico empleando un modelo trivariado para Chile* (Masters, Universidad Nacional de Colombia). Recuperado de <http://bdigital.unal.edu.co/46655/>

International Energy Agency. (2017). World Energy Balances 2017: Overview. Recuperado 20 de junio de 2018, de IEA webstore website: <https://webstore.iea.org/world-energy-balances-2017-overview>

Lozano Forero, S. (2018). *Propuesta de un índice SIPSA para el pronóstico de la inflación de Alimentos. Evidencia empírica*. (Especialización). Fundación Universitaria los Libertadores.

Posada, S. L., & Noguera, R. R. (2007). *Comparación de modelos matemáticos:*

Una aplicación en la evaluación de alimentos para animales. 8.

Rios, G., & Hurtador, C. (2008). *Series de Tiempo.* Universidad de Chile Facultad

de Ciencias Físicas y Matemáticas.

Rosales, R., Urrego, A., Sánchez, J. S., Correa, F., & Morales, C. (2008). *EJC 22:*

Metodología BOX – JENKINS. 5.

Unidad de Planeación Minero Energética UPME. (2016). *Informe de Gestión*

UPME 2015-2016. Recuperado de

www1.upme.gov.co/InformesGestion/Informe_gestion_2016.pdf

UPME. (2017). *Proyección regional de demanda de energía eléctrica y potencia*

máxima en Colombia. Recuperado de

http://www1.upme.gov.co/DemandaEnergetica/Proyeccion_demanda_regional_energia_electrica_2017.pdf

UPME. (2019, marzo 10). Recuperado 10 de marzo de 2019, de Unidad de

Planeación Minero Energética UPME website:

<http://www1.upme.gov.co/Entornoinstitucional/NuestraEntidad/Paginas/Quienes-Somos.aspx>

XM S.A. E.S.P. (2015, enero). *Informe Consolidado del Mercado Versión*

Liquidación TXR. Recuperado de

http://www.xm.com.co/Informes%20Mensuales%20de%20Anlisis%20del%20Mercado/00_Informe_Consolidado_TXR_01_2015.pdf

XM S.A. E.S.P. (2019). *Gestión Información Inteligente.* Recuperado 29 de julio de

2019, de <http://informacioninteligente10.xm.com.co/Pages/default.aspx>